Linear Algebra and its Applications 436 (2012) 1562–1571

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Some inequalities for positive linear maps

Rajendra Bhatia [a,*], Rajesh Sharma [b]

[a] Indian Statistical Institute, Delhi Centre, 7, S.J.S. Sansanwal Marg, New Delhi 110 016, India
[b] Department of Mathematics, Himachal Pradesh University, Shimla, India

ARTICLE INFO

Article history:
Received 25 March 2010
Accepted 16 September 2010
Available online 1 November 2010

Submitted by C. Davis

Dedicated to José Dias da Silva

MSC:
47A63
47C15
15A60
46L53

Keywords:
Positive unital linear map
Convex function
Jensen's inequality
Variance
Spread
Condition number

ABSTRACT

It has long been known that an analogue of Jensen's inequality holds for positive unital linear maps on matrix algebras provided that instead of ordinary convex functions one restricts to matrix convex functions. We show that this restriction is not necessary in the case of 2×2 matrices. A noncommutative analogue of the variance is studied, and a basic inequality, with several applications, is established.

1. Introduction

Let $\mathbb{M}(n)$ be the C^*-algebra of all $n \times n$ complex matrices and let $\Phi : \mathbb{M}(n) \to \mathbb{M}(k)$ be a positive unital linear map [1]. A fundamental inequality of Kadison [2] says that for every Hermitian matrix A

$$\Phi(A)^2 \leqslant \Phi(A^2). \tag{1}$$

This is a noncommutative analogue of the classical inequality

$$\mathbb{E}(X)^2 \leqslant \mathbb{E}(X^2), \tag{2}$$

* Corresponding author.
E-mail address: rbh@isid.ac.in (R. Bhatia), rajesh_-hpu_-math@yahoo.co.in (R. Sharma).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2010.09.038

where X is a random variable (a real-valued measurable function) on a probability space $(\Omega, \mathcal{F}, P)$ with finite expectation $\mathbb{E}X := \int XdP$. In classical analysis the inequality (2) is subsumed in the much more general *Jensen's inequality*. If the range of X is contained in the interval (a, b) and f is a convex function on (a, b), then

$$f(\mathbb{E}(X)) \leqslant \mathbb{E}(f(X)). \tag{3}$$

This leads to the natural problem of finding a general version of (1): if A is a Hermitian matrix whose spectrum is contained in (a, b) and f is a convex function on (a, b) then do we have

$$f(\Phi(A)) \leqslant \Phi(f(A)). \tag{4}$$

It was shown by Davis [3] that this is true when f is a *matrix convex* function [4] and Φ is *completely positive*. The latter restriction was removed by Choi [5] who showed that (4) remains valid for all positive unital linear maps Φ provided f is matrix convex. Very few convex functions are matrix convex. For example on the interval $(0, \infty)$ the function $f(t) = t^r$, $r > 0$, is convex for all $r \geqslant 1$ but matrix convex only for $1 \leqslant r \leqslant 2$. Choi gives an interesting example of a positive unital linear map Φ from $\mathbb{M}(3)$ to $\mathbb{M}(2)$ and a Hermitian matrix A for which the inequality $\Phi(A)^4 \leqslant \Phi(A^4)$ is false. (Let Φ be the compression map taking a 3×3 matrix A to the 2×2 matrix sitting in the top left corner of A and let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.)

In Choi's example $n = 3$ and $k = 2$. The value chosen for k is minimal. It is a well known fact that for every positive unital linear *functional* ("state") Φ on $\mathbb{M}(n)$ the inequality (4) is true for all convex functions f. See Problem IX.8.14 in [4].

In Section 2 we show that the value $n = 3$ in Choi's example is also minimal in the sense that the inequality (4) does hold for all convex functions f and all positive unital linear maps from $\mathbb{M}(2)$ into $\mathbb{M}(k)$. This is somewhat surprising as in most of these problems effects of noncommutativity already manifest themselves in the case of 2×2 matrices. We should remark that there has been considerable interest in recent years in 2×2 matrices because of quantum information and computing.

An inequality complementary to (1) was obtained by Bhatia and Davis [6]. This says that if the spectrum of a Hermitian matrix A is contained in the interval $[m, M]$, then

$$\Phi(A^2) - \Phi(A)^2 \leqslant \left(\frac{M - m}{2}\right)^2, \tag{5}$$

for every positive unital linear map Φ. The commutative case

$$\mathbb{E}(X^2) - \mathbb{E}(X)^2 \leqslant \left(\frac{M - m}{2}\right)^2, \tag{6}$$

has been known for long; see e.g. [7]. The quantity on the left hand side of (6) is called the *variance* of the random variable X. In analogy we call the expression on the left hand side of (5) the variance of A.

Kadison's inequality (1) was generalised by Choi [5,8] who showed that for every A in $\mathbb{M}(n)$ we have

$$\Phi(A)^*\Phi(A) \leqslant \Phi(A^*A), \tag{7}$$

provided Φ is 2-positive and unital. If A is normal, then this is true for all positive unital Φ. We seek an inequality complementary to this in the spirit of (5). Let

$$\Delta(A) = \inf_{z \in \mathbb{C}} \|A - z\| \tag{8}$$

be the distance of A from scalar matrices. We show that for all A in $\mathbb{M}(n)$ and for all positive unital Φ we have

$$\Phi(A^*A) - \Phi(A)^*\Phi(A) \leqslant \Delta(A)^2. \tag{9}$$

In the special case when A is normal

$$\Delta(A) = r_A, \tag{10}$$

where r_A is the radius of the smallest disk containing the spectrum of A. When A is Hermitian the inequality (9) reduces to (5).

As for many basic inequalities, the proof of (9) is short and simple but its consequences are many and surprisingly strong. We demonstrate some of these in Section 3, where we show the connection between this inequality and several old and new papers.

2. Jensen's Inequality for 2 × 2 matrices

We begin with a simple lemma on convex functions.

Lemma 2.1. *Let f be a real valued convex function on an interval containing (a, b). Then for $a \leqslant x \leqslant b$ we have*

$$f(x) \leqslant \frac{f(b) - f(a)}{b - a} x - \frac{af(b) - bf(a)}{b - a}. \tag{11}$$

Proof. Consider the function g defined as

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} x + \frac{af(b) - bf(a)}{b - a}.$$

Then $g(a) = g(b) = 0$, and for $x = (1 - t)a + tb$, $0 < t < 1$, we have

$$g(x) \leqslant (1 - t)f(a) + tf(b) - \frac{f(b) - f(a)}{b - a} ((1 - t)a + tb) + \frac{af(b) - bf(a)}{b - a}.$$

A little algebraic manipulation shows that the right hand side of the last inequality is equal to zero. This proves (11). □

Theorem 2.2. *Let $\Phi : \mathbb{M}(2) \to \mathbb{M}(k)$ be a positive unital linear map. Let f be a convex function on an open interval containing the eigenvalues of a Hermitian element A of $\mathbb{M}(2)$. Then*

$$f(\Phi(A)) \leqslant \Phi(f(A)). \tag{12}$$

Proof. If A is a scalar matrix the two sides of (12) are equal to $f(A)$. So assume A has distinct eigenvalues $\lambda_1 > \lambda_2$. Choose orthogonal projections P_1 and P_2 such that

$$P_1 + P_2 = I, \tag{13}$$

and

$$A = \lambda_1 P_1 + \lambda_2 P_2. \tag{14}$$

Then

$$f(A) = f(\lambda_1)P_1 + f(\lambda_2)P_2. \tag{15}$$

Apply the map Φ to the Eqs. (14) and (15) and then solve the two resulting equations to get

$$\Phi(P_1) = \frac{f(\lambda_2)\Phi(A) - \lambda_2\Phi(f(A))}{\lambda_1 f(\lambda_2) - \lambda_2 f(\lambda_1)}, \tag{16}$$

$$\Phi(P_2) = \frac{\lambda_1\Phi(f(A)) - f(\lambda_1)\Phi(A)}{\lambda_1 f(\lambda_2) - \lambda_2 f(\lambda_1)}, \tag{17}$$

assuming that $\lambda_1 f(\lambda_2) \neq \lambda_2 f(\lambda_1)$. From (13) we have

$$\Phi(P_1) + \Phi(P_2) = I. \tag{18}$$

From (16), (17) and (18) we obtain

$$\Phi(f(A)) = \frac{f(\lambda_1) - f(\lambda_2)}{\lambda_1 - \lambda_2}\Phi(A) - \frac{\lambda_2 f(\lambda_1) - \lambda_1 f(\lambda_2)}{\lambda_1 - \lambda_2}.$$

By Lemma 2.1 the right hand side of this inequality is greater than or equal to $f(\Phi(A))$, and we have proved the desired inequality (12). The restriction $\lambda_1 f(\lambda_2) \neq \lambda_2 f(\lambda_1)$ can be removed by a continuity argument, or by arguing the exceptional case separately. □

3. Bounds for the variance

Theorem 3.1. *Let $\Phi : \mathbb{M}(n) \to \mathbb{M}(k)$ be any positive unital linear map and let A be any element of $\mathbb{M}(n)$. Then*

$$\Phi(A^*A) - \Phi(A)^*\Phi(A) \leqslant \Delta(A)^2, \tag{19}$$

where $\Delta(A) = \inf_{z\in\mathbb{C}} \|A - z\|$.

Proof. Since $A^*A \leqslant \|A\|^2$ and Φ is positive and unital, we have

$$\Phi(A^*A) \leqslant \|A\|^2,$$

and therefore

$$\Phi(A^*A) - \Phi(A)^*\Phi(A) \leqslant \|A\|^2.$$

Using the properties of Φ it can easily be seen that the left hand side of this inequality is not changed when A is replaced by $A - z$. This leads to the inequality (19). □

The quantity $\Delta(A)$ has been studied in connection with several problems. In [9] Bhatia and Semrl showed that

$$2\Delta(A) = \max\left\{\|A - UAU^*\| : U \text{ unitary}\right\}, \tag{20}$$

the diameter of the *unitary orbit* of A. Using this they obtained a proof of a well known theorem of Stampfli [10]:

$$2\Delta(A) = \max_{\|X\|=1} \|AX - XA\|. \tag{21}$$

This quantity is the norm of the *derivation* $\delta_A(X) := AX - XA$. Another expression for $\Delta(A)$ was derived by Ando. This says

$$\Delta(A) = \max\left\{|\langle y, Ax\rangle| : \|x\| = \|y\| = 1 \text{ and } x \perp y\right\}. \tag{22}$$

A proof of this can be found in [9].

Let $x \in \mathbb{C}^n$ be a unit vector and let $\varphi(A) = \langle x, Ax\rangle$. Then φ is a positive unital linear functional on $\mathbb{M}(n)$. For this functional the quantity on the left hand side of (19) is

$$\text{var}_x(A) := \|Ax\|^2 - |\langle x, Ax\rangle|^2. \tag{23}$$

Let $[x]$ be the space spanned by the vector x and $[x]^\perp$ its orthogonal complement. With respect to the decomposition

$$\mathbb{C}^n = [x] \oplus [x]^\perp,$$

the vector Ax can be split as

$$Ax = \alpha x + \beta y,$$

where y is a unit vector orthogonal to x, $\alpha = \langle x, Ax\rangle$, $\beta = \langle y, Ax\rangle$ and $|\alpha|^2 + |\beta|^2 = \|Ax\|^2$. Combining this information with (22) and (23) we obtain the following theorem proved recently in [11] by Audenaert:

Theorem 3.2. *Let A be any $n \times n$ matrix. Then*

$$\max_{\|x\|=1} \left(\|Ax\|^2 - |\langle x, Ax\rangle|^2\right) = \Delta(A)^2. \tag{24}$$

Corollary 3.3. *Let r_A be the radius of the smallest disk in the complex plane that contains the spectrum of A. Then*

$$\max_{\|x\|=1} \left(\|Ax\|^2 - |\langle x, Ax\rangle|^2\right) \geqslant r_A^2. \tag{25}$$

When A is normal the two sides of (25) are equal.

Proof. By Schur's Theorem there exists a unitary matrix U such that $UAU^* = T = D + N$ where T is upper triangular, and D is the diagonal part of T. By well known properties of the norm

$$\|A - z\| = \|U(A - z)U^*\| = \|(D - z) + N\| \geqslant \|D - z\|.$$

The entries of D are the eigenvalues of A. Hence

$$\Delta(A) = \inf_z \ \|A - z\| \geqslant \inf_z \ \|D - z\| = r_A.$$

So the inequality (25) follows from (24). When A is normal, we have $N = 0$ and $\Delta(A) = r_A$. $\square$

The results in Corollary 3.3 have long been known for operators in Hilbert space. The statement about normal operators was proved by Björck and Thomée [12], that about arbitrary operators by Garske [13].

The linear functional $\varphi(A) = \langle x, Ax\rangle = \text{tr } xx^*A$ is one of a more general class. Let ρ be a positive semidefinite matrix of trace 1 (called a *density matrix* in the physics literature). Then $\varphi(A) = \text{tr } \rho A$ is a positive unital linear functional. If ρ has rank one, then $\rho = xx^*$ for some unit vector x. Choosing different density matrices ρ we obtain from the inequality (19) stronger versions of several known results. We illustrate this with a few examples.

Let e_j, $1 \leqslant j \leqslant n$, be the standard basis for $\mathbb{C}^n$. Let $\rho = e_je_j^*$. Then $\varphi(A) = a_{jj}$ and we obtain from (19)

$$\Delta(A)^2 \geqslant \max_j \sum_{k \neq j} |a_{kj}|^2. \tag{26}$$

Choosing $\rho = \frac{1}{n}I$ we get $\varphi(A) = \frac{1}{n}$ tr A, and then (19) leads to the inequality

$$\Delta(A)^2 \geqslant \frac{1}{n} \ \|A\|_F^2 - \frac{1}{n^2} \ |\text{tr } A|^2, \tag{27}$$

where $\|A\|_F$ is the *Frobenius norm* of A defined by the relation $\|A\|_F^2 = \text{ tr } A^*A$. In between these two examples is the following. Let $\mathcal{I}$ be any subset of $\{1, 2, \ldots, n\}$ and let ρ be $1/|\mathcal{I}|$ times the orthogonal projection on the span of the vectors $e_j, j \in \mathcal{I}$. Then

$$\varphi(A) = \frac{1}{|\mathcal{I}|} \sum_{j \in \mathcal{I}} a_{jj}. \tag{28}$$

With a little calculation we obtain from (19) the following.

Theorem 3.4. *Let $\mathcal{I}$ be any subset of $\{1, 2, \ldots, n\}$. Then*

$$\Delta(A)^2 \geqslant \frac{1}{|\mathcal{I}|} \sum_{\substack{i \in \mathcal{I} \\ j \neq i}} |a_{ji}|^2 + \frac{1}{|\mathcal{I}|^2} \sum_{\substack{i,j \in \mathcal{I} \\ i<j}} |a_{ii} - a_{jj}|^2. \tag{29}$$

The inequalities (26) and (27) are subsumed in this. Another special case when $\mathcal{I}$ is a set with two elements leads to:

Corollary 3.5. *For every matrix A*

$$\Delta(A)^2 \geqslant \frac{1}{2} \max_{i,j} \left\{ \sum_{k \neq i} |a_{ki}|^2 + \sum_{k \neq j} |a_{kj}|^2 + \frac{|a_{ii} - a_{jj}|^2}{2} \right\}. \tag{30}$$

Another interesting bound is obtained by taking $\rho = \frac{1}{n}E$, where E is the matrix with all entries equal to one. Then

$$\varphi(A) = \frac{1}{n} \sum_{i,j} a_{ij}. \tag{31}$$

For an easy comparison between various bounds let us introduce the quantity

$$\mathrm{dr}\ A = \sum_{i \neq j} a_{ij}. \tag{32}$$

Then $\varphi(A) = \frac{1}{n}(\mathrm{tr}\, A + \mathrm{dr}\, A)$. Let

$$v_1(A) = \frac{1}{n}\ \mathrm{tr}\ A^*A - \frac{1}{n^2}\ |\mathrm{tr}\, A|^2, \tag{33}$$

and

$$v_2(A) = \frac{1}{n}\ \mathrm{dr}\ A^*A - \frac{1}{n^2}\ |\mathrm{dr}\ A|^2. \tag{34}$$

Then $v_1(A) \geqslant 0$. The quantity $v_2(A)$ is always real and could be negative. (The matrix A^*A is positive semidefinite; the sum of all its entries and its trace are nonnegative numbers and $\mathrm{dr}\, A^*A$ is their difference.) With a little calculation we obtain from (19) the following.

Theorem 3.6. *For every matrix A*

$$\Delta(A)^2 \geqslant v_1(A) + v_2(A) - \frac{2\mathrm{Re}\,\mathrm{tr}\, A\ \mathrm{dr}\, A}{n^2}. \tag{35}$$

The quantity on the right hand side of (27) is $v_1(A)$. The bound (35) is better than (27) in many cases; for example when $\mathrm{dr}\, A = 0$ and $\mathrm{dr}\, A^*A \geqslant 0$. One such Hermitian matrix is

$$A = \begin{bmatrix} 1 & i & 2i \\ -i & 2 & 3i \\ -2i & -3i & 3 \end{bmatrix}.$$

Beginning with Mirsky [14] several authors have obtained bounds for the spread of a matrix, and ideas close to ours occur in their papers. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix A. The *spread* of A is defined as

$$\mathrm{spd}(A) = \max_{j \neq k}\ |\lambda_j - \lambda_k|. \tag{36}$$

If the λ_j are all real, then clearly $\mathrm{spd}\,(A) = 2r_A$ (where r_A is the radius of the smallest disk containing all λ_j). If the λ_j are arbitrary complex numbers, then a classical theorem of Jung (see [15, Chapter 16]) implies that

$$\mathrm{spd}\,(A) \geqslant \sqrt{3}\ \ r_A. \tag{37}$$

The example

$$A=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$$

in which the eigenvalues are the cube roots of 1 shows that the inequality (37) is best possible.

We have observed that $\Delta(A)=r_A$ when A is normal. So the lower bounds for $\Delta(A)$ lead to lower bounds for the spreads of Hermitian and normal matrices. Thus from (26) we obtain some results of Mirsky [16]. The inequality (27) leads to Theorem 3.2 of Barnes and Hoffman [17], and (30) to the main result of that paper. Theorem 3.4 restricted to Hermitian matrices gives a considerable strengthening of Theorem 6 of Jiang and Zhan [18]. Cognate results and comparisons between them can be found in [19]–[21].

A very ingenious use of inequality (5) has been made by Audenaert in [11]. He has used it to provide a marvelous proof of a commutator estimate obtained earlier by Böttcher and Wenzel [22]. This says that for any two $n\times n$ matrices A and X we have

$$\|AX-XA\|_2\leqslant\sqrt{2}\,\|A\|_2\ \|X\|_2. \tag{38}$$

When A is normal there is an easy proof for this. It was observed in [23] that in this special case (A normal) we have for every unitarily invariant norm $|||\cdot|||$

$$|||AX-XA|||\leqslant\sqrt{2}\,\mathrm{spd}\,(A)\,|||X|||.$$

The authors, not aware of Jung's theorem, applied an inequality weaker than (37) with $\sqrt{2}$ in place of $\sqrt{3}$. With (37) in hand this can be strengthened to

$$|||AX-XA|||\leqslant\frac{2}{\sqrt{3}}\,\mathrm{spd}\,(A)\ \ |||A|||. \tag{39}$$

This inequality is sharp. If we choose

$$A=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix},\quad X=\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix},$$

then $\mathrm{spd}\,(A)=\sqrt{3}$, $\|AX-XA\|_1=2$, and $\|X\|_1=1$. Here $||\cdot||_1$ stands for the trace norm (the sum of all singular values). By a duality argument, the inequality (39) is sharp for the operator norm $||\cdot||$ as well.

An inequality stronger than (5) was obtained in [6]. This says that

$$\Phi(A^2)-\Phi(A)^2\leqslant(M-\Phi(A))(\Phi(A)-m), \tag{40}$$

for every positive unital linear map Φ and every Hermitian matrix A whose spectrum is in the interval $[m,M]$. Using this inequality we obtain lower bounds on the condition number of a positive definite matrix.

Theorem 3.7. *Let φ be a positive unital linear functional on $\mathbb{M}(n)$ and let A be a positive definite matrix whose eigenvalues are in the interval $[m,M]$. Let*

$$\alpha=\varphi(A),\quad \beta^2=\varphi(A^2)-\varphi(A)^2.$$

Then

$$\frac{M}{m} \geqslant \left(\frac{\beta}{\alpha} + \sqrt{1 + \left(\frac{\beta}{\alpha}\right)^2}\right)^2. \tag{41}$$

Proof. From the inequality (40) we can obtain

$$\frac{M}{m} \geqslant \frac{\beta^2 + \alpha^2 - m\alpha}{m(\alpha - m)}. \tag{42}$$

Let $f(m)$ be the expression on the right hand side of (42). It can be seen that

$$f'(m) = \frac{-\alpha(m-a)(m-b)}{m^2(\alpha - m)^2},$$

where

$$a = \frac{\beta^2 + \alpha^2 - \sqrt{(\beta^2 + \alpha^2)\beta^2}}{\alpha},$$

and

$$b = \frac{\beta^2 + \alpha^2 + \sqrt{(\beta^2 + \alpha^2)\beta^2}}{\alpha}.$$

It follows that $f(m)$ has its minimum at $m = a$. A little calculation then shows that

$$f(m) \geqslant \left(\frac{\alpha}{\sqrt{\beta^2 + \alpha^2} - \beta}\right)^2 = \left(\frac{\sqrt{\beta^2 + \alpha^2} + \beta}{\alpha}\right)^2.$$

The last expression is equal to the right hand side of (41). □

Special choices of φ in the theorem above lead to interesting lower bounds for M/m easily computable from the entries of A. Choosing $\varphi(A) = \frac{1}{n} \operatorname{tr} A$, one gets from (41)

$$\frac{M}{m} \geqslant \left(\frac{\sqrt{n \operatorname{tr} A^2} + \sqrt{n \operatorname{tr} A^2 - (\operatorname{tr} A)^2}}{\operatorname{tr} A}\right)^2. \tag{43}$$

Choosing $\varphi(A)$ as in (28) we see that

$$\frac{M}{m} \geqslant \max_{i,j} \left(\frac{\beta(i,j)}{\alpha(i,j)} + \sqrt{1 + \left(\frac{\beta(i,j)}{\alpha(i,j)}\right)^2}\right)^2, \tag{44}$$

where

$$\alpha(i,j) = \frac{a_{ii} + a_{jj}}{2},$$

and

$$\beta(i,j)^2 = \frac{1}{2}\left\{\sum_{k \neq i} |a_{ik}|^2 + \sum_{k \neq j} |a_{jk}|^2 + \frac{(a_{ii} - a_{jj})^2}{2}\right\}.$$

Example 1. We borrow the following example from [19]. Let

$$A = \begin{bmatrix} 4 & 0 & 2 & 3 \\ 0 & 5 & 0 & 1 \\ 2 & 0 & 6 & 0 \\ 3 & 1 & 0 & 7 \end{bmatrix}.$$

Then from the bound in Corollary 2.3 of [19] we get $M/m \geqslant 2.4953$. From our inequality (43) we obtain $M/m \geqslant 2.7238$ and from *(44)* $M/m \geqslant 3.5353$.

4. Remarks

(1) The argument of Theorem 3.1 shows also that

$$\Phi(AA^*) - \Phi(A)^* \, \Phi(A) \leqslant \Delta(A)^2. \tag{45}$$

Using this instead of (19) we can obtain different versions of some of the subsequent inequalities.

(2) It might be useful to record the commutative version of some of the statements in Section 3. Let X be a complex valued random variable. Define the *variance* of X as

$$\mathrm{var}(X) = \mathbb{E}|X|^2 - |\mathbb{E}(X)|^2.$$

Let r_X be the radius of the smallest disk containing the range of X, and $\mathrm{spd}(X)$ the maximum distance between any two points in this range. Then

$$\mathrm{var}(X) \leqslant r_X^2 \leqslant \frac{1}{3}\, \mathrm{spd}(X)^2.$$

(3) Let P be an orthogonal projection of rank $k \leqslant n/2$. Then $\Phi_P(A) = PAP$ is called the *compression* of A onto the range of P. If we take P to be the projection onto the span of the first k basis vectors $e_1, \ldots, e_k$, then $\Phi_P(A)$ is the top left $k \times k$ block in the matrix A. Let A be a Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. A well-known inequality of Bloomfield and Watson [24], then says that

$$\mathrm{tr}\, \left(\Phi_P(A^2) - \Phi_P(A)^2\right) \leqslant \frac{1}{4} \sum_{j=1}^{k} \left(\lambda_j - \lambda_{n-j+1}\right)^2 . \tag{46}$$

This inequality is both subtler and stronger than the inequality

$$\frac{1}{k}\, \mathrm{tr}\, \left(\Phi_P(A^2) - \Phi_P(A)^2\right) \leqslant \frac{1}{4}\, (\lambda_1 - \lambda_n)^2. \tag{47}$$

The functional $\varphi(A) = \frac{1}{k}\, \mathrm{tr}\, \Phi_P(A)$ is positive and unital. So from (5) one obtains

$$\frac{1}{k}\, \mathrm{tr}\, \Phi_P(A^2) - \frac{1}{k^2}\, (\mathrm{tr}\ \Phi_P(A))^2 \leqslant \frac{1}{4}\, (\lambda_1 - \lambda_n)^2. \tag{48}$$

The left hand side of (48) is not smaller than that of (47). So the inequality (48) is not included in (47).

Many very striking and powerful inequalities generalising those of Bloomfield and Watson were obtained by Ando [25]. A very readable and comprehensive exposition with an extensive bibliography can be found in [26].

(4) The left hand side of (9) has been interpreted as the variance of A. Likewise the *covariance* between two matrices A and B can be defined as

$$\text{cov}\,(A, B) = \Phi(A^*B) - \Phi(A)^*\Phi(B). \tag{49}$$

A variance–covariance inequality in this context was formulated and proved in [27]. See also [1].

Acknowledgements

The first author is supported by a J.C. Bose National Fellowship. The second author is supported by the UGC-SAP. He thanks the Indian Statistical Institute for a visit in January 2010 when this work was done.

References

[1] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2007.
[2] R.V. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras, Ann. Math. 56 (1952) 494–503.
[3] C. Davis, A Schwarz inequality for convex operator functions, Proc. Amer. Math. Soc. 8 (1957) 42–44.
[4] R. Bhatia, Matrix Analysis, Springer, 1996.
[5] M.D. Choi, A Schwarz inequality for positive linear maps on C^*-algebras, Illinois J. Math. 18 (1974) 565–574.
[6] R. Bhatia, C. Davis, A better bound on the variance, Amer. Math. Monthly 107 (2000) 353–357.
[7] T. Popoviciu, Sur les équations algébriques ayant toutes leurs racines réelles, Mathematica 9 (1935) 129–145.
[8] M.D. Choi, Some assorted inequalities for positive linear maps on C^*-algebras, J. Operator Theory 4 (1980) 271–285.
[9] R. Bhatia, P. Semrl, Orthogonality of matrices and some distance problems, Linear Algebra Appl. 287 (1999) 77–85.
[10] J.G. Stampfli, The norm of a derivation, Pacific J. Math. 33 (1970) 737–747.
[11] K.M.R. Audenaert, Variance bounds with an application to norm bounds for commutators, Linear Algebra Appl. 432 (2010) 1126–1143.
[12] G. Björck, V. Thomée, A property of bounded normal operators in Hilbert space, Ark. Mat. 4 (1963) 551–555.
[13] G. Garske, An equality concerning the smallest disc that contains the spectrum of an operator, Proc. Amer. Math. Soc. 78 (1980) 529–532.
[14] L. Mirsky, The spread of a matrix, Mathematika 3 (1956) 127–130.
[15] H. Rademacher, O. Toeplitz, The Enjoyment of Mathematics, Dover, 1990.
[16] L. Mirsky, Inequalities for normal and hermitian matrices, Duke Math. J. 24 (1957) 591–598.
[17] E.R. Barnes, A.J. Hoffman, Bounds for the spectrum of normal matrices, Linear Algebra Appl. 201 (1994) 79–90.
[18] E. Jiang, X. Zhan, Lower bounds for the spread of a Hermitian matrix, Linear Algebra Appl. 256 (1997) 153–163.
[19] H. Wolkowicz, G.P.H. Styan, Bounds for eigenvalues using traces, Linear Algebra Appl. 29 (1980) 471–506.
[20] C.R. Johnson, R. Kumar, H. Wolkowicz, Lower bounds for the spread of a matrix, Linear Algebra Appl. 71 (1985) 161–173.
[21] K.J. Merikoski, R. Kumar, Characterization and lower bounds for the spread of a normal matrix, Linear Algebra Appl. 364 (2003) 13–31.
[22] A. Böttcher, D.Wenzel, How big can the commutator of two matrices be and how big is it typically? Linear Algebra Appl. 403 (2005) 216–228.
[23] R. Bhatia, F. Kittaneh, Commutators, pinchings, and spectral variation, Oper. Matrices 2 (2008) 143–151.
[24] P. Bloomfield, G.S. Watson, The inefficiency of least squares, Biometrika 62 (1975) 124–128.
[25] T. Ando, Bloomfield–Watson–Knott type inequalities for eigenvalues, Taiwanese J. Math. 5 (2001) 443–469.
[26] S.W. Drury, S. Liu, C.-Y. Lu, S. Puntanen, G.P.H. Styan, Some comments on several matrix inequalities with applications to canonical correlations: historical background and recent developments, Sankhyā Ser. A 64 (2002) 453–507.
[27] R. Bhatia, C. Davis, More operator versions of the Schwarz inequality, Commun. Math. Phys. 215 (2000) 239–244.